

# Méthodes numériques : optimisation.

## Préparation au partiel de mars 2015 : éléments de correction

Amic Frouvelle

4 mars 2015

### 1 Cas pathologique de la méthode de Newton

On considère la fonction  $x \mapsto e^{x^4}$  de  $\mathbb{R}$  dans  $\mathbb{R}$ . On cherche à minimiser  $f$  en approchant un zéro de  $f'$  par la méthode de Newton.

1. On a  $f'(x) = 4x^3e^{x^4}$ , puis  $f''(x) = (12x^2 + 16x^6)e^{x^4}$ . On obtient donc

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} = x_k - \frac{4x_k^3}{12x_k^2 + 16x_k^6} = x_k \left(1 - \frac{1}{3 + 4x_k^4}\right).$$

2. On a  $1 - \frac{1}{3+x_k^4}$  qui appartient à  $[\frac{2}{3}, 1[$ , donc la suite  $(|x_k|)_{k \in \mathbb{N}}$  est non nulle et strictement décroissante dès que  $x_0 \neq 0$ . Elle converge donc vers une limite  $x_*$ , qui satisfait  $x_* = x_* \left(1 - \frac{1}{3+x_*^4}\right)$ , ce qui n'est possible que si  $x_* = 0$ .

La suite  $(x_k)$  converge donc vers 0, et on a

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - 0|}{|x_k - 0|} = \lim_{k \rightarrow \infty} \left(1 - \frac{1}{3 + x_k^4}\right) = \frac{2}{3} < 1.$$

La convergence est donc linéaire et le taux de convergence linéaire est  $\frac{2}{3}$ .

3. Dans des hypothèses raisonnables, on s'attend à ce que la méthode de Newton fournisse des itérés qui convergent quadratiquement. Ici ce n'est pas le cas, même lorsque l'on démarre très proche de l'unique minimum. L'hypothèse cruciale qui n'est pas satisfaite est qu'au point de minimum local  $x_*$ , il faut qu'on ait  $f''(x_*) > 0$ , de sorte que  $f$  est bien approximée par une parabole, pour avoir une convergence quadratique lorsqu'on démarre suffisamment proche du minimum local (il faut également suffisamment de régularité, une condition suffisante est par exemple  $f$  de classe  $C^3$ ). Ici, on a  $f''(0) = 0$  : le point critique est dégénéré.

4. Les courbes sont tracées en échelle logarithmique pour les ordonnées (on a tracé  $|x_k - 0| = x_k$  en fonction de  $k$ ). On voit bien que la pente asymptotique de toutes ces courbes est la même, ce qui correspond à ce qu'on attend de suites qui convergent linéairement avec le même taux de convergence  $\frac{2}{3}$ . Cependant, ce taux n'est qu'une limite, et lorsque l'on démarre avec des conditions initiales trop grandes ( $x_0 = 2$  ou  $x_0 = 4$ ), alors la décroissance initiale est très lente et il faut un certain nombre d'itérations pour pouvoir « voir » la convergence linéaire (plus de 40 itérations si  $x_0 = 4$ ).

Cela correspond au fait qu'initialement, le rapport entre  $x_{k+1}$  et  $x_k$  donné par  $1 - \frac{1}{3+4x_k^4}$  est très proche de 1 : si  $x_0 = 4$ , alors  $\frac{1}{3+4 \cdot 4^4} = \frac{1}{3+2^{10}} \approx 10^{-3}$ , ce qui veut dire qu'au départ, les  $x_k$  ne sont diminués que de 0,1% environ. . .

## 2 Précision d'une méthode de recherche de minimum

1. Comme on divise par  $\varepsilon$  ou  $2\varepsilon$  au dénominateur dans les formules de différences finies, l'erreur qu'on fait au numérateur entre  $f$  et  $\hat{f}$  induit des erreurs numériques de l'ordre de  $\frac{\eta}{\varepsilon}$ , ce qui explose si  $\varepsilon$  est petit comparativement à  $\eta$ .

La formule des différences finies centrées est dans ce cas plus efficace, car elle permet d'atteindre une meilleure précision (d'ordre  $\eta^{\frac{2}{3}} \ll \sqrt{\eta}$ , puisqu'on a supposé  $\eta \ll 1$ ), et ceci pour le même coût : 2 évaluations de fonction. Ceci est dû au fait que l'erreur théorique de la méthode des différences finies centrées est de l'ordre de  $\varepsilon^2$ , plutôt que d'ordre  $\varepsilon$  (en faisant la différence finie centrée, on annule le terme d'ordre 2 dans le développement de Taylor). Du coup, on peut se permettre de prendre un  $\varepsilon$  plus grand tout en gardant au moins la même précision, ce qui donne une erreur numérique bien plus petite. On peut donc augmenter la précision jusqu'à ce que l'erreur numérique soit du même ordre.

Pour les différences finies à droite, on obtient la meilleure précision lorsque les erreurs sont du même ordre :  $\varepsilon \approx \frac{\eta}{\varepsilon}$  ce qui donne  $\varepsilon$  de l'ordre de  $\sqrt{\eta}$ , et une précision d'ordre également  $\sqrt{\eta}$ . Et pour les différences finies centrées, c'est lorsque  $\varepsilon^2 \approx \frac{\eta}{\varepsilon}$  c'est à dire  $\varepsilon$  de l'ordre de  $\eta^{\frac{1}{3}}$ , et la précision est donc d'ordre  $\eta^{\frac{2}{3}}$ .

2. Supposons que  $\hat{f}'_d(x) > 0$ , alors on a

$$f'(x) - \hat{f}'_d(x) \geq -|f'(x) - \hat{f}'_d(x)| \geq -C_d \|f\|_{\infty} \sqrt{\eta}.$$

En prenant par exemple  $C_0 > C_d \|f\|_{\infty}$ , on obtient que si  $\hat{f}'_d(x) \geq C_0 \sqrt{\eta}$ , alors

$$f'(x) \geq (C_0 - C_d \|f\|_{\infty}) \sqrt{\eta} > 0.$$

Si on suppose  $\hat{f}'_d(x) < 0$ , et si  $\hat{f}'_d \leq -C_0 \sqrt{\eta}$ , on obtient

$$f'(x) \leq \hat{f}'_d + |f'(x) - \hat{f}'_d(x)| \leq (C_d \|f\|_{\infty} - C_0) \sqrt{\eta} < 0.$$

Donc dans les deux cas on a  $f'(x)$  qui est strictement du même signe que  $\hat{f}'_d(x)$ .

La méthode de dichotomie appliquée à  $\hat{f}'_d$  fournit donc exactement les mêmes itérées que si on l'avait appliquée à  $f'$  tant que l'on a  $|\hat{f}'_d(x)| \geq C_0 \sqrt{\eta}$ .

On va d'abord essayer d'obtenir une estimation sur  $|\hat{f}'_d(x)|$  pour tout  $x$ . Supposons par exemple que  $x > x_*$ , par le théorème des accroissements finis appliqué à  $f'$  en  $x_*$ , on sait qu'il existe un  $c \in [x_*, x]$  tel que  $f'(x) = f'(x_*) + (x - x_*)f''(c) \geq h(x - x_*)$ , puisque  $f'(x_*) = 0$ . De même si  $x < x_*$ , alors  $f'(x) \leq -h(x - x_*)$ , et donc dans tous les cas on a  $|f'(x)| \geq h|x - x_*|$ .

Donc on a

$$|\hat{f}'_d(x)| \geq |f'(x)| - |\hat{f}'_d(x) - f'(x)| \geq h|x - x_*| - C_d \|f\|_{\infty} \sqrt{\eta},$$

Et donc si  $h|x - x_*| \geq (C_0 + C_d \|f\|_{\infty}) \sqrt{\eta}$ , c'est à dire si  $|x - x_*| \geq C_* \sqrt{\eta}$  (où  $C_* = \frac{C_0 + C_d \|f\|_{\infty}}{h}$ ), alors  $|\hat{f}'_d(x)| \geq C_0 \sqrt{\eta}$  et donc  $\hat{f}'_d(x)$  et  $f'(x)$  sont strictement du même signe.

On aurait pu montrer directement ce fait-là, et c'est suffisant pour la suite, mais l'énoncé était bancal sur ce point-là, mea culpa.

On part donc du fait suivant :

- Il existe une constante  $C_*$  de l'ordre de grandeur de un, telle que si on a  $|x - x_*| \geq C_* \sqrt{\eta}$ , alors  $\hat{f}'_d(x)$  et  $f'(x)$  sont strictement du même signe.

On sait que la méthode de dichotomie fournit toujours deux suites convergentes  $a_n$  et  $b_n$  qui sont des suites adjacentes, et on note  $x_{\infty}$  leur limite. On sait également qu'on a toujours  $\hat{f}'_d(a_n)$  et  $\hat{f}'_d(b_n)$  de signes opposés.

Supposons par exemple qu'on ait  $a_n > x_*$ , alors comme  $b_n > a_n$  on a  $f'(a_n) > 0$  et  $f'(b_n)$ , on ne peut pas avoir  $|a_n - x_*| \geq C_* \sqrt{\eta}$ , sinon on aurait  $\hat{f}'_d(a_n) > 0$  et également  $|b_n - x_*| > |a_n - x_*| \geq C_* \sqrt{\eta}$  et donc  $\hat{f}'_d(b_n) > 0$ .

Donc si  $a_n > x_*$ , alors  $|a_n - x_*| < C_* \sqrt{\eta}$ . De même on montre que si  $b_n < x_*$ , alors  $|b_n - x_*| < C_* \sqrt{\eta}$ . Si on choisit par exemple d'arrêter l'algorithme dès que  $|b_n - a_n| \leq \sqrt{\eta}$ , on note  $n_0$  la première étape où cela arrive, et on pose  $\hat{x}_* = \frac{a_{n_0} + b_{n_0}}{2}$ . Alors

— soit  $a_{n_0} \leq x_* \leq b_{n_0}$ , et dans ce cas  $|\hat{x}_* - x_*| \leq \frac{\sqrt{\eta}}{2}$ .

— soit  $a_{n_0} > x_*$ , dans ce cas  $|\hat{x}_* - x_*| \leq |\hat{x}_* - a_{n_0}| + |a_{n_0} - x_*| < \frac{\sqrt{\eta}}{2} + C_* \sqrt{\eta}$ .

— soit  $b_{n_0} < x_*$ , et de même  $|\hat{x}_* - x_*| < (\frac{1}{2} + C_*) \sqrt{\eta}$ .

Si on prend  $C_1 = \frac{1}{2} + C_*$ , on obtient bien que

$$|\hat{x}_* - x_*| \leq C_1 \sqrt{\eta}.$$

On peut même montrer que le critère d'arrêt est satisfait dès que  $\frac{b-a}{2^n} \leq \sqrt{\eta}$ , vu que la taille de l'intervalle  $[a_n, b_n]$  diminue d'un facteur deux à chaque itération. Cela revient à  $n \geq \frac{1}{\ln 2} \ln(\frac{b-a}{\sqrt{\eta}})$ . Le coût total de la méthode en terme du nombre d'évaluations de fonctions est de l'ordre de  $\frac{2}{\ln 2} \ln(\frac{b-a}{\sqrt{\eta}})$ , puisqu'il y a 2 évaluations pour approximer la dérivée au point milieu.

- Il y avait une erreur d'énoncé, c'était évidemment pour le cas où on utilisait  $\hat{f}'_c$ . Dans ce cas-là, tous les raisonnements précédents fonctionnent exactement de la même manière en remplaçant  $\sqrt{\eta}$  par  $\eta^{\frac{2}{3}}$ . En effet on n'a jamais utilisé la spécificité des différences finies à droite, mais simplement l'estimation  $|f'(x) - \hat{f}'_d(x)| \leq C_d \sqrt{\eta} \|f\|_\infty$ . On peut donc avoir une précision sur  $x_*$  de l'ordre de  $\eta^{\frac{2}{3}}$ , ce qui est donc mieux.
- Pour voir quelle est la précision qu'on obtient sur l'approximation de  $f(x_*)$ , on écrit le développement de Taylor à l'ordre deux en  $x_*$ , appliqué en  $\hat{x}_*$  :

$$f(\hat{x}_*) = f(x_*) + \frac{1}{2} f''(x_*) |\hat{x}_* - x_*|^2 + o(|\hat{x}_* - x_*|^2) = f(x_*) + f''(x_*) \frac{C_1^2}{2} \eta + o(\eta).$$

On obtient donc que

$$\begin{aligned} |\hat{f}(\hat{x}_*) - f(x_*)| &\leq |\hat{f}(\hat{x}_*) - f(\hat{x}_*)| + |\hat{f}(\hat{x}_*) - f(x_*)| \\ &\leq \eta \|f\|_\infty + f''(x_*) \frac{C_1^2}{2} \eta + o(\eta). \end{aligned}$$

Si on veut avoir un peu plus de détail sur le terme  $o(\eta)$ , il faut faire une hypothèse supplémentaire sur  $f$ , par exemple que  $f''(x) \leq H$  sur  $[a, b]$  avec  $H$  une constante de l'ordre de un, de telle sorte qu'on remplace le  $\frac{1}{2} f''(x_*) |\hat{x}_* - x_*|^2 + o(|\hat{x}_* - x_*|^2)$  dans la formule de Taylor par une formule exacte, par exemple  $\frac{1}{2} f''(c) |\hat{x}_* - x_*|^2$  où  $c$  est un réel entre  $x_*$  et  $\hat{x}_*$ . On obtient alors

$$|\hat{f}(\hat{x}_*) - f(x_*)| \leq (\|f\|_\infty + \frac{1}{2} H C_1^2) \eta.$$

On a donc une précision de l'ordre de  $\eta$  sur l'approximation de la valeur de  $f(x_*)$ .

Si on avait utilisé les différences centrées, l'estimation de Taylor aurait donné un meilleur résultat (de l'ordre de  $\eta^{\frac{4}{3}}$ ), mais l'autre terme restant serait toujours  $\eta \|f_\infty\|$ , donc on n'améliore pas la précision totale cette fois-ci. Les deux méthodes sont donc aussi précises l'une que l'autre.

- Si on ne suppose plus que  $f'' > h \geq 0$  sur  $[a, b]$ , en particulier si  $f''(x_*) = 0$ , alors l'estimation que l'on avait pour dire que  $f'(x)$  et  $\hat{f}'_d(x)$  sont du même signe ne fonctionne plus : on peut avoir  $|x - x_*| \geq C \sqrt{\eta}$  et  $|f'(x)|$  d'un ordre de grandeur plus petit, de telle sorte que même en sachant que  $|f'(x) - \hat{f}'_d(x)|$  est de l'ordre de  $\sqrt{\eta}$ , cela ne nous aide pas à savoir le signe de  $\hat{f}'(x)$ , car  $|f'(x)|$  est trop petit. Par exemple si  $f(x) = x^4 + 1$ , on a alors  $x_* = 0$  et  $f'(x) = 4x^3$  qui est de l'ordre de grandeur de  $\eta^{\frac{3}{2}}$  (ce qui est bien plus petit que  $\sqrt{\eta}$ ) si  $|x - 0|$  est de l'ordre de grandeur de  $\sqrt{\eta}$ .

## 3 Étude de cas en dimension 2

### 3.1 Étude de la fonction

1. On a  $F(r, \theta + 2k\pi) = F(r, \theta)$ , donc si  $X = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}$ ,  $r$  est uniquement déterminé par  $r = \sqrt{x^2 + y^2}$ , et  $\theta$  est déterminé à  $2k\pi$  près, donc la valeur du terme de droite ne change pas quelque soit le choix que l'on fait pour  $\theta$ , et ne dépend donc bien que de  $x$  et de  $y$ .
2. La fonction est bien  $C^1$  par composition sur  $\{r < 1\}$  et sur  $\{r > 1\}$ . Il suffit de montrer que les dérivées partielles se prolongent par continuité pour obtenir que  $f$  est  $C^1$  sur  $\mathbb{R}^2$ .

On les calcule en écrivant  $F(r, \theta) = f(r \cos \theta, r \sin \theta)$ , que l'on dérive par rapport à  $r$  et  $\theta$  avec la règle de la chaîne :

$$\begin{aligned}\partial_r F(r, \theta) &= \cos \theta \partial_x f(r \cos \theta, r \sin \theta) + \sin \theta \partial_y f(r \cos \theta, r \sin \theta) \\ \partial_\theta F(r, \theta) &= -r \sin \theta \partial_x f(r \cos \theta, r \sin \theta) + r \cos \theta \partial_y f(r \cos \theta, r \sin \theta).\end{aligned}$$

On résout alors en multipliant la première ligne par  $\cos \theta$  et la deuxième par  $-\frac{\sin \theta}{r}$  et en sommant pour obtenir  $\partial_x f$ , puis en multipliant la première ligne par  $\sin \theta$  et la deuxième par  $\frac{\cos \theta}{r}$  et sommant pour obtenir  $\partial_y f$ . On obtient

$$\begin{aligned}\partial_x f(r \cos \theta, r \sin \theta) &= \partial_r F(r, \theta) \cos \theta - \frac{1}{r} \partial_\theta F(r, \theta) \sin \theta \\ \partial_y f(r \cos \theta, r \sin \theta) &= \partial_r F(r, \theta) \sin \theta + \frac{1}{r} \partial_\theta F(r, \theta) \cos \theta,\end{aligned}$$

ce qui est exactement la formule demandée pour  $r > 1$ .

On calcule alors

$$\begin{aligned}g(r, \theta) = \partial_r F(r, \theta) &= 2(r-1)(2 - \cos(4 \ln(r-1) + \theta)) + \frac{4}{r-1}(r-1)^2 \sin(4 \ln(r-1) + \theta) \\ &= 2(r-1)(2 - \cos(4 \ln(r-1) + \theta) + 2 \sin(4 \ln(r-1) + \theta)) \\ &= 2(r-1)(2 + \sqrt{5}(-\sin \theta_0 \cos(4 \ln(r-1) + \theta) + \cos \theta_0 \sin(4 \ln(r-1) + \theta))) \\ &= 2(r-1)(2 + \sqrt{5} \sin(4 \ln(r-1) + \theta - \theta_0)),\end{aligned}$$

ce qui donne la formule attendue. Enfin le calcul de  $h(r, \theta) = \frac{1}{r} \partial_\theta F(r, \theta)$  est direct et donne la formule attendue.

Il ne reste plus qu'à montrer que ces dérivées partielles sont continues au voisinage de  $r = 1$ . En fait on a  $|g(r, \theta)| \leq 2(r-1)(2 + \sqrt{5})$  lorsque  $r > 1$  et  $|h(r, \theta)| \leq (r-1)^2$ , donc  $\|\nabla f(X)\| \leq 2(r-1)(2 + \sqrt{5}) + (r-1)^2$  puisque  $\|\mathbf{u}(\theta)\| = \|\mathbf{v}(\theta)\| = 1$ . On obtient donc que  $\nabla f(X)$  converge vers  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  lorsque  $\|X\|$  tend vers 1 tout en étant strictement supérieure à 1. Les dérivées partielles se prolongent donc par continuité au voisinage du cercle unité, et donc la fonction est bien  $C^1$  sur  $\mathbb{R}^2$ .

3. Si  $X$  est un point critique avec  $r > 1$ , alors puisque  $\mathbf{u}(\theta)$  et  $\mathbf{v}(\theta)$  forment une base (orthogonale) de  $\mathbb{R}^2$ , on doit avoir  $g(r, \theta) = 0$  et  $h(r, \theta) = 0$ . On obtient donc d'après cette deuxième condition que  $4 \ln(r-1) + \theta = k\pi$  avec  $k \in \mathbb{Z}$ , et que donc  $g(r, \theta) = 2(r-1)(2 + \sqrt{5} \sin(k\pi - \theta_0))$ , c'est-à-dire  $g(r, \theta) = 2(r-1)(2 \pm \sqrt{5} \sin \theta_0) = 2(r-1)(2 \pm 1)$ , qui est différent de 0 dès que  $r > 1$ . Il n'y donc pas de point critique autre que tous ceux de la boule unité.

Comme d'après la définition de  $f$ , il est évident que  $f$  est positive sur  $\mathbb{R}^2$ , on obtient que ces points sont des points de minimum global.

4. La base  $(\mathbf{u}(\theta), \mathbf{v}(\theta))$  étant orthonormale, on a  $\|\nabla f(X)\|^2 = g(r, \theta)^2 + h(r, \theta)^2$ . forment une base (orthogonale) de  $\mathbb{R}^2$ . Comme  $|g(r, \theta)| \leq 2(r-1)(2 + \sqrt{5})$  lorsque  $r > 1$  et  $|h(r, \theta)| \leq r-1$ , on en conclut que  $\|\nabla f(X)\|^2 \leq (r-1)^2(4(2 + \sqrt{5})^2 + 1)$ . Comme on a  $(2 + \sqrt{5})^2 = 4 + 5 + 4\sqrt{5} \leq 9 + 12$  (car  $\sqrt{5} \leq 3$ ), on obtient que  $(4(2 + \sqrt{5})^2 + 1) \leq 85 \leq 100$ , donc on a bien  $\|\nabla f(X)\|^2 \leq 100(r-1)^2$ .

### 3.2 Méthode de descente de gradient à pas fixe

1. On a  $f(X_0) = 1$  (on a  $r = 2$  et  $\cos \theta = 1$ ). Et on a  $f(X) \geq (r-1)^2(2-1) = (r-1)^2$ , donc si  $r > 2$ , alors  $f(X) > 1$ . Donc  $S_0$  est bien inclus dans la boule de centre 0 et de rayon 2. Comme  $f$  est continue  $S_0 = f^{-1}([0, 1])$  est fermé, donc compact puisque borné. Puisque le pas  $\alpha$  vérifie  $\alpha < \frac{2}{L}$ , que la fonction est bornée inférieurement et que son gradient est  $L$ -lipschitzien sur  $\mathbb{R}^2$ , il est bien  $L$ -lipschitzien sur  $S_0$ , et le théorème du cours s'applique, donnant que  $f(X_k)$  est décroissante, converge vers une valeur critique (d'après le corollaire où on suppose  $S_0$  compact), donc vers 0, et que  $\|\nabla f(X_k)\|$  tend vers 0 lorsque  $k \rightarrow \infty$ .
2. On a, si  $r_k > 1$

$$r_{k+1} = \|X_{k+1}\| = \|X_k - \alpha \nabla f(X_k)\| \geq \|X_k\| - \alpha \|\nabla f(X_k)\| \geq r_k - 10\alpha(r_k - 1),$$

d'après la question 4. de l'étude de la fonction. Et donc on a

$$(r_{k+1} - 1) \geq r_k - 1 - 10\alpha(r_k - 1) \geq (r_k - 1)(1 - 10\alpha) > 0.$$

Ce qui donne par récurrence que  $r_k > 1$  pour tout  $k$  (puisque  $r_0 = 2$ ), et donne la partie gauche de l'estimation voulue. Pour l'autre partie, on écrit simplement que

$$r_{k+1} - 1 = \|X_k - \alpha \nabla f(X_k)\| - 1 \leq \|X_k\| + \alpha \|\nabla f(X_k)\| - 1 \leq r_k - 1 + 10\alpha(r_k - 1) \leq (r_k - 1)(1 + 10\alpha).$$

Comme on a vu que  $f(X) \geq (r-1)^2$ , comme  $f(X_k) \rightarrow 0$ , on obtient que  $r_k \rightarrow 1$ .

3. Attention à l'ambiguïté de la notation  $\theta_0$  (qu'on a déjà utilisé dans l'étude de la fonction). J'aurais dû écrire  $\theta_*$  à la place de  $\theta_0$  pour ne pas avoir de problème de notation. On a

$$X_{k+1} \cdot X_k = r_{k+1} r_k \mathbf{u}(\theta_{k+1}) \cdot \mathbf{u}(\theta_k) = r_{k+1} r_k (\cos \theta_{k+1} \cos \theta_k + \sin \theta_{k+1} \sin \theta_k) = r_{k+1} r_k \cos(\theta_{k+1} - \theta_k).$$

On obtient donc que  $\cos(\theta_{k+1} - \theta_k) = \frac{X_{k+1} \cdot X_k}{r_{k+1} r_k} = \frac{X_{k+1} \cdot \mathbf{u}(\theta_k)}{r_{k+1}}$ . On a en fait  $X_{k+1} = X_k - \alpha \nabla f(X_k) = (r_k - \alpha g(r_k, \theta_k)) \mathbf{u}(\theta_k) - \alpha h(r_k, \theta_k) \mathbf{v}(\theta_k)$ , de sorte que lorsqu'on prend le produit scalaire avec  $\mathbf{u}(\theta_k)$ , il ne reste que le premier terme. De plus, on obtient la norme de  $X_{k+1}$  qui correspond au dénominateur de la formule voulue puisque  $\mathbf{u}(\theta_k), \mathbf{v}(\theta_k)$  est une base orthonormale.

On a donc

$$\sin^2(\theta_{k+1} - \theta_k) = 1 - \cos^2(\theta_{k+1} - \theta_k) = \frac{\alpha^2 h(r_k, \theta_k)^2}{r_{k+1}^2}.$$

Quitte à rajouter un nombre entier de fois  $2\pi$ , on peut supposer que  $\theta_{k+1}$  a été défini de sorte que  $|\theta_{k+1} - \theta_k|$  soit le plus petit possible. Comme  $r_k \rightarrow 1$ ,  $g(r_k, \theta_k) \rightarrow 0$  et  $h(r_k, \theta_k) \rightarrow 0$  lorsque  $k \rightarrow \infty$ , on obtient que  $\cos(\theta_{k+1} - \theta_k) \rightarrow 1$ , donc à partir d'un certain rang  $\cos(\theta_{k+1} - \theta_k) \geq 0$ , et donc  $\theta_{k+1} - \theta_k \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ . On peut donc appliquer l'estimation (le facteur est  $\frac{\pi}{2}$  et non pas  $\frac{2}{\pi}$ , il y avait une coquille dans l'énoncé) :

$$|\theta_{k+1} - \theta_k| \leq \frac{\pi}{2} |\sin(\theta_{k+1} - \theta_k)| = \frac{\pi}{2r_{k+1}} \alpha |h(r_k, \theta_k)| \leq \frac{\pi}{2r_{k+1}} \alpha (r_k - 1)^2.$$

Comme  $r_{k+1} \rightarrow 1$ , il existe un rang à partir duquel  $\frac{\pi}{2r_{k+1}} \leq 2$  et on obtient donc l'estimation voulue (il y avait une coquille, il manquait le facteur 2 et il y avait un  $a$  à la place du 1 dans l'énoncé) :

$$|\theta_{k+1} - \theta_k| \leq 2\alpha (r_k - 1)^2.$$

4. Cette question était difficile et aurait mérité une indication. Mea culpa de nouveau.

L'idée est de trouver un intervalle tel que si  $t_k$  est dans cet intervalle, la fonction  $g(r_k, \theta_k)$  est négative, de sorte que  $r_{k+1} \geq r_k$ . Si on arrive même à être en dessous d'une constante négative, on va pouvoir montrer que  $r_k$  augmente suffisamment pour que  $4 \ln(r_k - 1)$  augmente plus que la

diminution possible de  $\theta_k$ . Notons  $\theta_*$  ce qui avait été noté  $\theta_0$  dans l'étude de fonction pour éviter la confusion.

Pour  $t_* = -\frac{\pi}{2} + \theta_*$ , on a  $2 + \sqrt{5} \sin(t_* - \theta_*) = 2 - \sqrt{5} < 0$ , alors par continuité il existe  $\delta > 0$  et un intervalle  $[a, b]$  contenant  $t_0$ , tel que pour tout  $t \in [a, b]$ , on ait  $2 + \sqrt{5} \sin(t - \theta_0) \leq -\delta$ . Remarquons que  $a$  et  $b$  ne dépendent pas de  $\alpha$ , et que si on prend  $t \in [a + 2n\pi, b + 2n\pi]$ , on a le même résultat.

Si on a  $t_k \in [a, b]$ , alors  $g(r_k, \theta_k) = 2(r_k - 1)(2 + \sqrt{5} \sin(t_k - \theta_0)) \leq -2\delta(r_k - 1)$ .

Et donc  $r_{k+1} = \sqrt{(r_k - \alpha g(r_k, \theta_k))^2 + \alpha^2 h(r_k, \theta_k)^2} \geq r_k + 2\alpha\delta(r_k - 1)$ . On en conclut que  $r_{k+1} - 1 \geq (r_k - 1)(1 + 2\alpha\delta)$ .

On a alors

$$\begin{aligned} t_{k+1} - t_k &= 4(\ln(r_{k+1} - 1) - \ln(r_k - 1)) + \theta_{k+1} - \theta_k \\ &\geq 4 \ln\left(\frac{r_{k+1} - 1}{r_k - 1}\right) - |\theta_{k+1} - \theta_k| \\ &\geq 4 \ln(1 + 2\alpha\delta) - 2\alpha(r_k - 1)^2. \end{aligned}$$

Cette quantité est positive dès que  $k$  est assez grand, puisque  $r_k \rightarrow 1$ .

5. On écrit que

$$\begin{aligned} |t_{k+1} - t_k| &= \left| 4 \ln\left(\frac{r_{k+1} - 1}{r_k - 1}\right) + \theta_{k+1} - \theta_k \right| \\ &\leq 4 \ln(1 + 10\alpha) + 2\alpha(r_k - 1)^2, \end{aligned}$$

d'après la question 2. On obtient donc que si  $\alpha$  est suffisamment petit, on peut avoir  $4 \ln(1 + 10\alpha) < b - a$ . Et donc pour ce choix de  $\alpha$  il existe un rang à partir duquel  $|t_{k+1} - t_k| < b - a$ .

Soit  $k_0$  un tel rang satisfaisant également la conclusion de la question 5. Alors il existe  $n_0$  tel que pour tout  $k \leq k_0$ ,  $t_k \geq b - 2n_0\pi$  (il suffit de prendre  $n_0$  suffisamment grand). Montrons alors que pour tout  $k$ , on a  $t_k \geq a - 2n_0\pi$ .

Si on suppose par l'absurde qu'il existe un  $k$  tel que  $t_k < a - 2n_0\pi$ , on note  $k_1$  le plus petit des entiers satisfaisant ça. On a alors  $k_1 > k_0$ , et donc  $|t_{k_1-1} - t_{k_1}| > b - a$  ce qui donne que  $t_{k_1-1} \in [a - 2n_0\pi, b - 2n_0\pi]$ . De la même manière que précédemment, cet intervalle satisfait la conclusion de la question 5. Ceci nous donne que  $t_{k_1} \geq t_{k_1-1} \geq a - 2n_0\pi$  et c'est donc une contradiction.

On a donc bien  $t_k \geq a - 2n_0\pi$  pour tout  $k \geq 0$ .

6. Comme  $\ln(r_k - 1) \rightarrow -\infty$  et que  $t_k$  est bornée inférieurement, la seule possibilité est que  $\theta_k \rightarrow +\infty$ . Si on se fixe un point de la forme  $\begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$  du cercle unité, on pose  $k = \varphi(n)$  le premier indice tel que  $\theta_k \geq t + 2n\pi$ . On a  $|\theta_{\varphi(n)} - t - 2n\pi| = \theta_{\varphi(n)} - t - 2n\pi < \theta_{\varphi(n)} - \theta_{\varphi(n)-1} \rightarrow 0$ , on obtient donc  $\theta_{\varphi(n)} - 2n\pi \rightarrow t$  et  $r_{\varphi(n)} \rightarrow 1$ , autrement dit  $X_{\varphi(n)} \rightarrow \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ .